DOCUMENT RESUME

ED 320 574 IR 053 169

AUTHOR White, Marilyn Domas
TITLE The Arizona Clinical Interview Rating Scale: Its Applicability and Adaptability for the Evaluation of Pre-Search Reference Interviews.
PUB DATE Feb 90
NOTE 16p.
PUB TYPE Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Field Tests; *Interpersonal Communication; *Interviews; *Questioning Techniques; *Rating Scales; *Reference Services; Reliability; Validity
IDENTIFIERS *Query Negotiation (Library Science); Validity Research

ABSTRACT

This paper discusses the applicability and adaptability for evaluating reference interviews in library science of the Arizona Clinical Interview Rating Scale (ACIRS), a process-oriented scale originally developed in 1976 at the Arizona College of Medicine to evaluate the interview performance of medical students. The analysis surveys the content and reliability and validity data and discusses the appropriateness of the criteria and the presentation for reference interviews. It is concluded that the ACIRS is an effective interview scale emphasizing process-related criteria that can be modified fairly easily for application to pre-search interviews. (9 references) (GL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED320574

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

THE ARIZONA CLINICAL INTERVIEW RATING SCALE: Its Applicability and Adaptability for the Evaluation of Pre-Search Reference Interviews

Marilyn Domas White, Associate Professor
College of Library and Information Services
University of Maryland
College Park, Maryland 20742

February, 1990

IR053169

BEST COPY AVAILABLE

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Marilyn Domas White

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ABSTRACT

This paper discusses the applicability and adaptability of the Arizona Clinical Interview Rating Scale [ACIRS] for evaluating reference interviews in library science. This scale was developed in 1976 by Stillman and others at the University of Arizona College of Medicine to evaluate the interview performance of medical students. Other researchers have used it in similar studies. The scale emphasizes process-related criteria, not content, and has been used with a separate content checklist. The analysis surveys the content and reliability and validity data and discusses the appropriateness of the criteria and the presentation for reference interviews. This analysis was done as part of a project to develop an evaluation form for the reference interview preceding online searches which can be used for performance evaluation in job settings. The ACIRS is not included in the report.

THE ARIZONA CLINICAL INTERVIEW RATING SCALE: Its Applicability and Adaptability for the Evaluation of Pre-Search Reference Interviews

1. Introduction

As part of a project to develop an evaluation form for the reference interview preceding online searches which can be used for performance evaluation in job settings, several interviewing scales used in other fields were analyzed for their general applicability and adaptability to reference interviews in library science. This paper discusses briefly the Arizona Clinical Interview Rating Scale [ACIRS], developed by Stillman and others at the University of Arizona College of Medicine in 1976. It is included as an appendix in Stillman (1976). It was used initially to evaluate the performance of medical students after a formal program in interviewing surrogate mothers with evaluative feedback instituted as part of a Pediatrics Clerkship. (Stillman et al., 1976; Stillman et al., 1977a; Stillman et al., 1977b) Other researchers have used it in similar studies, occasionally modifying the scale. (Carroll et al., 1981)

2. Description of the Scale

2.1 Content and Organization

The following outline indicates the skills evaluated on Arizona Clinical Interview Rating Scale and the polar positions

of the scales used:

1. Organization
    - 1.1 All parts in proper sequence/missing parts
    - 1.2 Focuses on one area at a time/irrelevant areas
2. Timeline
    - 2.1 Logical progression/Haphazard, unrelated progression
3. Transitional statements
    - 3.1 Transitional statements between sections/No explanations leaving uncertainty about purpose of questioning
4. Questioning skills
    - 4.1 Forced choice questions/leading questions
    - 4.2 No unnecessary digressions, smooth flow/ interruption of continuity
    - 4.3 Repetition only for clarification/Frequent repetition to obtain information already provided
    - 4.4 Consistent use of summary statements to verify or clarify/Never summarizing
    - 4.5 Understandable language in questions/Use of jargon and unexplained technical terms
5. Documentation of data
    - 5.1 Verification and specificity/Accepting information at face value
6. Rapport

6.1 Eye contact/No eye contact

6.2 Attentive, no interruptions/Detachment, interruption

6.3 Sensitivity to non-illness related concerns and indepth exploration/Unalert, avoiding possible involvement

6.4 Social reinforcement, feedback/Little support, emphasis on negative

6.5 Encouraged additional questions/Allowed no opportunity to bring up additional questions

6.6 Summarized all pertinent information/No attempt to summarize

Parts 1 and 2 both address structure and order in the interview. The ideal interview, according to the ACIRS, would cover all appropriate content areas (introduction, chief complaint and history of present illness, past medical history, social and family history, and review of systems) in this order. The interviewer would concentrate on one topic at a time and confined his questions to that topic. In subtopics, especially the major one about the client's problem, he would follow a logical, usually chronological, order.

Part 3 indirectly impacts on structure and order in its consideration of transition statements and emphasizes use, in a good interview, of transitional statements to lead the client through the sections and to clarify the importance or need for

specific questions. Covered in Parts 4.4 and 6.6 are summary statements for individual subsections, and a summary statement for the interview as a whole, which also influence perception of organization, and could be considered to facilitate transition. (5)

Parts 4 and 5 focus on question tactics. The scale optimizes a smooth, flowing interview with no interruptions, the interviewer's asking for specific information to validate the client's generalities, using closed, forced choice questions to move efficiently through a subarea which has high, but predictable, information content, no extraneous questions, repetition only for clarification or validation, reliance on standard English, not jargon or technical language in questions, and summary statements for each subportion of the interview.

Part 6 concerns rapport and emphasizes maintaining eye contact, attentive, empathetic response to the client, sensitivity to subtle, perhaps seemingly unrelated, factors underlying the client's problem and discussion of these factors in adequate depth, providing warm, supportive feedback, and active solicitation of additional information during closure. Closure is characterized by a summary statement containing information gathered during the interview.

2.2 Format

The ACIRS is a summative measuring instrument, listing sixteen interviewing skills in six areas. Each skill is defined with a five-point scale, with anchoring statements for excellent,

average, and poor performance. The scales are established so that the optimal score is 5 in all cases. Total points which can be obtained are 80 points.

2.3 Reliability and Validity

Formal testing was done to assess the reliability and validity of the scale before its original use. The sixteen skills included in the instrument were identified as skills discriminating good and bad interviews after observation of interviews done by expert physician interviewers. On the final version of the scale, interjudge reliability by paraprofessional judges involved as subjects in interviews was .87; intrajudge reliability involving evaluation of the same interviews at two weeks' interval was .85 and .90. Scores for a group with a pediatric clerkship in which interviewing skills were significantly higher than those for a group which had no interview training. Several additional studies indicate that interviewing instruction does result in higher ACIRS scores. Insignificant correlations between scores on the ACIRS and various subtests of the Medical College Admissions Test indicate that the ACIRS does not measure scholastic or medical aptitude. Coefficients of internal consistency (.79 and .80 for groups of 36 and 60 students respectively) are high, indicating the ACIRS measures essentially one trait. (Stillman et al., 1977c)

On the basis of their studies, the developers conclude that "the ACIR Scale should be useful in establishing reliable evaluation of interviewing skills for a variety of case

histories. In addition to providing for outcome evaluation (e.g. grading decisions) the ACIR Scale can furnish information useful for formative evaluation (e.g. monitoring of student progress to recommend further instruction) ... ." (Stillman et al., 1977c)

A separate study by Swanson and others presents conflicting results. They assessed the interjudge reliability, intercase reliability, and construct validity of three instruments on a comparative basis. They applied the ACIRS, an interaction analysis form, and a instrument consisting of two content-related checklists and counts of barriers to communication to 93 physician/patient encounters. Estimates of interjudge reliability estimates (intraclass correlations based on Winer (1971)) were derived from 24 interviews. They are single rater reliability coefficients with the differences between raters considered measurement errors. ACIRS had the lowest reliabilities, "reflecting the subjectivity involved in rating scales generally." Interjudge reliability was .77 for all items (.87 for the first nine dealing with questioning skills; .66 for the last five concerned with rapport).

Looking at comparability of ratings across two interviews by the same interviewer, intercase reliability with all instruments was "markedly lower" than the interjudge reliabilities. "ACIRS' intercase reliabilities are sufficiently low that mean scores derived from five patient encounters would have a reliabliity of only 0.64, ignoring interrater reliability."

In assessing construct validity, ACIRS and the information

analysis instrument did not reveal expected differences in interview quality across years of residency or programs with different training. Results for the checklists/counts instrument were mixed. The researchers commented, "on the whole ... evidence for validity of the measures was lacking." (Swanson, et al., 1981)

3. Evaluation

3.1 General Guidelines

In considering a scale or evaluation form's adaptability to the reference interview preceding online searches, several factors are considered: semantics/syntax, context, and appropriateness of criteria. The last is the most substantive. In addition, the evaluation also considers any constraints on applying the scale to interviews observed in real time or recorded in different media. In many cases, the scales are somewhat field-specific. Instead of using generic terminology, such as interviewer and interviewee, they refer instead to the positions in the field which correspond to these classifications. Frequently these are simply cosmetic differences applied to instruments which have broader applicability. It is relatively easy to apply more generic terms or to adjust them to include positions within this field. Although not necessary, such specificity, especially in examples, enhances the reliability of coding using the scale.

A more critical problem is the use of contextual examples which are field-specific. Often these point to real differences

underlying interviewing across fields. The importance of this aspect varies across the scales. They can be relatively minor if the examples constitute a negligible aspect of the scale and they can be eliminated without markedly affecting the scale or if there are reasonably appropriate counterparts in library science.

The most significant question for the purpose of this analysis is the appropriateness of the criteria being evaluated. In this respect, two considerations have to be made: first, the appropriateness of the criteria included, and secondly, the completeness of the criteria, even within circumscribed parameters, such as process.

3.2 Evaluation of the ACIRS

In describing modifications or problems, it is presumed that higher order problems subsume lower order problems. The scale incorporates 48 statements, reflecting measurements along certain criteria. Fifteen of these can be used as is.

3.2.1 Semantic Modifications

Only semantic modification would have to occur in 21 of the remaining 33 statements but could be done fairly easily in almost all cases simply by substituting more general terms or information science equivalents.

3.2.2 Contextual Difficulties

Three statements would require more substantial modification to incorporate nuances in the library science environment.

3.2.3 Appropriateness of Criteria

This interview scale does not claim to judge content in an

interview, although at least one portion directly addresses the idea of clearly established content and an ideal sequence for covering it. In the original research using this interview, content was addressed separately via the use of a detailed content list which was simply checked to indicate that the information had been obtained. The percentage of content identified was the measure of content quality. Instead the interview form's emphasis is on process-related variables.

Many of these are excellent process-related criteria and several address characteristics of a good reference interview already discussed in library and information science literature.

Several elements seem questionable in their applicability to the pre-search reference interview:

- prescribing an ideal sequence for progressing through prescribed topical areas or an overall sequence of topics.
- prescribing a chronological approach within certain topical areas
- considering forced choice and leading questions as opposite ends of a scale, with the result of optimizing forced choice questions, albeit within parameters, but the parameters are not clear.
- calling for exploration in depth of areas of expressed concern which may seem not immediately relevant to the problem under discussion.

While it may be possible to identify topical areas which should be included in the pre-search interview because of the

nature of the tasks associated with search strategy, specifying sequencing seems unnecessarily rigid. It seems more reasonable to stipulate the need for a perceptibly logical sequence, without indicating what that arrangement should be. This would allow for flexibility in responding to a range of problems and/or questions. Similarly, prescribing a logical, but not necessarily chronological, order in addressing topics within subareas seems more desirable than prescribing chronological arrangement. In all fairness, the interviewer would lose relatively few points if his sequencing were not the prescribed sequence.

In these criteria, the objection is more to how the scale is developed rather than to the objectives of the scale. But, in the last two, the objection is more to the actual criteria or to perceived problems which would make judgments on it unreliable.

The forced choice/leading question range is the only type-of-question variable considered and seems out of place when little attention is given to prescribing an adequate mix of open and closed questions, which logically should proceed focusing on one type of closed question in one type of subarea. The scale does specify a context, i.e. in areas where the students is required to deal with a large amount of potential information (e.g., history of present illness and review of systems) where this should be optimized.

In connection with rapport, the scale which emphasizes exploring in sufficient depth expressed concerns which do not seem to have obvious relevance to the question. Assessing when

following up such a lead is admirable or merely digressive and counter-productive is difficult. It seems possible that this scale would conflict with judgments on other questions which seem to prioritize efficiency in information-gathering.

3.2.4 Media Sensitivity

In the Stillman studies, the scale was used in real-time situations and with videotaped interviews. But, because only one criterion directly measures non-verbal behavior, and this is one of six considered in the rapport area, it may be possible to apply the ACIRS to audio-taped interviews by eliminating that criterion.

4. Conclusion

The ACIRS is a good interview scale emphasizing process-related criteria which can be modified fairly easily for application to pre-search interviews. Stillman and associates claim to avoid content and suggest use of the rating scale with a content checklist, but the organization area establishes certain content areas which seem task specific. In addition, the scale addresses the concept of optimal mix of types of questions only indirectly by assessing use of a particular type of closed question in a particular kind of situation.

Bibliography

Carroll, J. G., Schwartz, M. W., and Ludwig, S. (1981). "An Evaluation of Simulated Patients as Instructors - Implications for Teaching Medical Interview Skills." **Journal of Medical Education** 56, 522-524.

Gardner, M. E., and Burpeau-Digregorio, M. Y. (1985). "Objective Assessment of Pharmacy Students Interviewing Skills." **American Journal of Pharmaceutical Education** 49, 137-44.

Stillman, P. L., Burpeau-Digregorio, M. Y., Nicholson, G. I., Sabers, D. L., and Stillman, A. E. (1983). "Six Years of Experience Using Patient Instructors to Teach Interviewing Skills." **Journal of Medical Education**, 58, 941-46.

Stillman, P. L., Morrow, G., and Sabers, D. L. (1977b). "Evaluation of Interviewing Skills for Pediatric Housestaff." Dallas, TX. (Meeting of the American Board of Pediatrics.)

Stillman, P. L., Sabers, D. L., and Redfield, D. L. (1977a). "Use of Trained Mothers to Teach Interviewing Skills to First-Year Medical Students: A Follow-up Study." **Pediatrics** 60, 165-69.

Stillman, Paula L., Brown, D., Redfield, D., and Sabers, D. (1977c). "Construct Validation of the Arizona Clinical Interview Rating Scale." **Educational & Psychological Measurement** 37, 1031-1038.

Stillman, Paula L., Sabers, Darrell L., and Redfield, D. L. (1976). "The Use of Paraprofessionals to Teach Interviewing Skills." **Pediatrics** 57, 769-74.

Swanson, David B., Mayewski, Raymond J., Norsen, Lisa, Baran, Gertrude, and Mushlin, Alvan I. (1981). "A Psychometric Study of Measures of Medical Interviewing Skills." (Proceedings of the 20th Annual Conference on Research in Medical Education, 3-8).

Winer, B. (1971). **Statistical Principles in Experimental Design.** McGraw Hill, New York.